(51) International Patent Classification⁷: **C12N 15/00**

(21) International Application Number: PCT/US00/22613

(22) International Filing Date: 11 August 2000 (11.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/151,224   26 August 1999 (26.08.1999)   US
60/172,128   17 December 1999 (17.12.1999)   US

(71) Applicant: CALGENE LLC [US/US]; 1920 Fifth Street, Davis, CA 95616 (US).

(72) Inventor: FILLATTI, Joanne, J.; 36757 Russell Blvd., Davis, CA 95616 (US).

(74) Agent: EAGLE, Alissa, M.; Calgene LLC, 1920 Fifth Street, Davis, CA 95616 (US).

(81) Designated States (national): AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW.

(84) Designated States (regional): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published:
— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NUCLEIC ACID SEQUENCES AND METHODS OF USE FOR THE PRODUCTION OF PLANTS WITH MODIFIED POLYUNSATURATED FATTY ACIDS

(57) Abstract: By this invention, novel nucleic acid sequences are provided, wherein said nucleic acid sequence is a genomic sequence of a plant desaturase encoding sequence. Also provided in the present invention are the promoter and intron sequences of the desaturase genomic sequences. Furthermore, recombinant DNA constructs employing the polynucleotide sequences are provided. The instant invention also provides methods for the modification of fatty acid compositions in host plant cells.

NUCLEIC ACID SEQUENCES AND METHODS OF USE FOR THE

PRODUCTION OF PLANTS WITH MODIFIED POLYUNSATURATED FATTY

ACIDS

5

INTRODUCTION

This application claims priority to U.S. Provisional Application Serial No.

60/151,224 filed August 26, 1999 and U.S. Provisional Application Serial No. 60/172,128

filed December 17, 1999.

10

Technical Field

The present invention is directed to nucleic acid sequences and constructs, and

methods related thereto.

15   ## Background

Plant oils are used in a variety of applications. Novel vegetable oils compositions

and/or improved means to obtain oils compositions, from biosynthetic or natural plant

sources, are needed. Depending upon the intended oil use, various different fatty acid

compositions are desired.

20   One means postulated to obtain such oils and/or modified fatty acid compositions

is through the genetic engineering of plants. However, it is necessary to identify the

appropriate nucleic acid sequences which are capable of producing the desired phenotypic

result, regulatory regions capable of directing the correct application of such sequences,

and the like.

Higher plants appear to synthesize fatty acids via a common metabolic pathway

(fatty acid synthetase pathway). In developing seeds, where fatty acids are attached to

glycerol backbones, forming triglycerides, for storage as a source of energy for further

germination, the FAS pathway is located in the proplastids. The first committed step is

5   the formation of acetyl-ACP (acyl carrier protein) from acetyl-CoA and ACP catalyzed by

the enzyme, acetyl-CoA:ACP transacylase (ATA). Elongation of acetyl-ACP to 16- and

18- carbon fatty acids involves the cyclical action of the following sequence of reactions:

condensation with a two-carbon unit from malonyl-ACP to form a ß-ketoacyl-ACP (ß-

ketoacyl-ACP synthase), reduction of the keto-function to an alcohol (ß-ketoacyl-ACP

10   reductase), dehydration to form an enoyl-ACP (ß-hydroxyacyl-ACP dehydrase), and

finally reduction of the enoyl-ACP to form the elongated saturated acyl-ACP (enoyl-ACP

reductase). ß-ketoacyl-ACP synthase I, catalyzes elongation up to palmitoyl-ACP

(C16:0), whereas ß-ketoacyl-ACP synthase II catalyzes the final elongation to stearoyl-

ACP (C18:0). Common plant unsaturated fatty acids, such as oleic, linoleic and a-

15   linolenic acids found in storage triglycerides, originate from the desaturation of stearoyl-

ACP to form oleoyl-ACP (C18:1) in a reaction catalyzed by a soluble plastid Δ-9

desaturase (also often referred to as "stearoyl-ACP desaturase"). Molecular oxygen is

required for desaturation in which reduced ferredoxin serves as an electron co-donor.

Additional desaturation is effected sequentially by the actions of membrane bound Δ-12

20   desaturase and Δ-15 desaturase. These "desaturases" thus create mono- or

polyunsaturated fatty acids respectively.

Obtaining nucleic acid sequences capable of producing a phenotypic result in

FAS, desaturation and/or incorporation of fatty acids into a glycerol backbone to produce

an oil is subject to various obstacles including but not limited to the identification of

25   metabolic factors of interest, choice and characterization of an enzyme source with useful

kinetic properties, purification of the protein of interest to a level which will allow for its

amino acid sequencing, utilizing amino acid sequence data to obtain a nucleic acid

sequence capable of use as a probe to retrieve the desired DNA sequence, and the

preparation of constructs, transformation and analysis of the resulting plants.

5   Thus, additional nucleic acid targets and methods for modifying fatty acid

compositions are needed. In particular, constructs and methods to produce a variety of

ranges of different fatty acid compositions are needed.

SUMMARY OF THE INVENTION

10

The present invention is generally directed to genomic desaturase polynucleotides,

and in particular to genomic desaturase polynucleotides which encode enzymes that

catalyze the insertion of a double bond into a fatty acyl moiety at the twelfth (Δ12

desaturase or fad2) or fifteenth (Δ15 desaturase or fad3) carbon position in a fatty acyl

15   chain as counted from the carboxyl terminus. Further, the present invention provides

isolated non-coding regions of such genomic polynucleotide sequences, particularly

including the introns, and promoter regions. Specific oligonucleotides are provided which

include partial or complete sequences which are derived from Δ12 and Δ15 desaturase

promoter and intron sequences. Although the sequences disclosed herein are obtained

20   from soybean plants, it is contemplated that additional sequences can be derived from

intron and promoter regions of desaturase genomic polynucleotide sequences which are

homologous or have identity to the soybean desaturase sequences. Such additional

desaturase sequences can be obtained using standard methods described below from a

variety of plant sources, in particular oilseed crops.

It is also an aspect of the present invention to provide recombinant DNA

constructs which can be used for the modification of the fatty acid composition in a plant

and in particular, to modify the transcription or transcription and translation (expression)

of desaturase genes or proteins, such as Δ12 and Δ15 desaturase. The invention is

5   particularly directed to DNA constructs which include sequences which are derived from

the intron or promoter regions of a genomic clone wherein said sequences are in a sense

or antisense orientation in a DNA construct. These DNA constructs are then used to

transform or transfect host cells to produce plants with modified levels of fatty acids,

particularly modified levels of oleic, linoleic and linolenic acid. It is particularly

10   contemplated to provide constructs and methods for down regulating Δ12 and Δ15

desaturase gene expression, so as to increase the levels of oleic acid and to decrease the

levels of linoleic acid and linolenic acid. It is particularly contemplated to alter the fatty

acid composition in seed tissue of oilseed crops.

The modified plant cells, plants, seeds and oils obtained by the expression of the

15   Δ12 and Δ15 desaturase polynucleotides are also considered part of the invention.

Further, it is contemplated to produce oil compositions with specific relative levels of

each fatty acid. One preferred embodiment comprises at least about 80 - 85% oleic acid,

no more than about 1 - 2% linoleic acid, and no more than about 1 - 3% linolenic acid;

and a second preferred embodiment comprising at least about 50-75% oleic acid, at least

20   about 10 - 30% linoleic acid, and no more than about 3% linolenic acid.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to genomic desaturase sequences, particularly the

isolated non-coding sequences from genomic fatty acid desaturase nucleic acid sequences

25   from host cell sources. A desaturase sequence of this invention includes any nucleic acid

4

genomic sequence, including all non-coding regions, encoding amino acids from a source,

such as a protein, polypeptide or peptide, obtainable from a cell source, which is capable

of catalyzing the insertion of a double bond into a fatty acyl moiety in a plant host cell,

i.e., in vivo, or in a plant cell-like environment, i.e. in vitro. As will be described in more

5   detail below, specific genomic polynucleotide sequences encoding enzymes which add

double bonds at the twelfth (Δ12 desaturase) and fifteenth (Δ15 desaturase) carbon

positions in a fatty acyl chain as counted from the carboxyl terminus are provided. In

addition, provided herein are specific non-coding regions of such genomic sequences.

The term "non-coding" refers to sequences of polynucleotides that do not encode

10   part or all of an expressed protein. Non-coding sequences include but are not limited to

introns, promoter regions, and 5' untranslated regions.

The term "intron" as used herein refers to the normal sense of the term as meaning

a segment of polynucleotides, usually DNA, that does not encode part or all of an

expressed protein.

15   The term "exon" as used herein refers to the normal sense of the term as meaning

a segment of polynucleotides, usually DNA, that encodes part or all of an expressed

protein.

Thus, the term "intron" refers to gene regions that are transcribed into RNA

molecules, but which are spliced out of the RNA before the RNA is translated into a

20   protein. As contrasted to the term "exon" which refers to gene regions that are transcribed

into RNA and subsequently translated into proteins.

As set forth in detail in the sequence listing and the examples, genomic Δ12

desaturase and Δ15 desaturase sequences and intron and promoter regions obtained from

such sequences are provided herein. In particular, two Δ12 desaturase genomic clones

were identified and are set forth in SEQ ID NOs:1 and 23. A single Δ15 desaturase

genomic clone was identified and is set forth in SEQ ID NO:3. A single intron region

was obtained from each of the Δ12 desaturase genomic clones with the sequences

provided in SEQ ID NOs:2 and 24, respectively. The promoter region from each of the

5   Δ12 desaturase genomic clones are respectively included in SEQ ID NO:1 (base pairs 1 -

1094) and SEQ ID NO:23 (base pairs 1 - 1704). The Δ15 desaturase included seven

introns in the coding region (set forth as SEQ ID NOs:4, 5, 6, 7, 8, 25 and 26). In

addition, preliminary results suggest that there is an additional intron within the 5'

untranslated region.

10   Although the sequences described herein are obtained from soybean, it is

contemplated that intron and promoter regions can be obtained from desaturase genomic

polynucleotide sequences which are homologous or have identity to the soybean

desaturase sequences. In particular, sequences can be obtained from other plant sources

and particularly from oilseed crops. Such genomic sequences can be obtained using

15   standard methods, certain of which are described below.

The sequences of the present invention can be used to modify the fatty acid

composition in a plant (see Example 3 and Table I). In particular, it is shown that sense

and antisense suppression can be used to obtain broad ranges in the levels of oleic,

linoleic and linolenic acid. In particular, it is shown that levels of oleic acid can range

20   from about 26 to 80 %, levels of linoleic acid can range from about 2.97 to 49.92 % and

levels of linolenic acid can range from about 3.38 to 8.81%. However, these are merely

representative of the broad range that be can achieved. Moreover, it is contemplated that

combinations of the sequences could be used to achieve additional fatty acid

compositions. Certain compositions are preferred based on the intended use of the oil.

One preferred composition includes at least about 50 - 75% oleic acid, at least about 10 -

30% linoleic acid and no more than about 3% linolenic acid. A particularly preferred

embodiment includes at least about 60 - 70% oleic acid, at least about 15 - 20% linoleic

acid and no more than about 3% linolenic acid.

5   Although the examples set forth herein utilize sense or antisense suppression to

downregulate the gene of interest, it is contemplated that other means of modifying gene

expression can be used. In particular, it is contemplated that gene expression can be down

regulated using DNA binding proteins which can be designed to specifically bind to the

non-coding regions identified herein or that ribozymes can be designed to cleave such

10   non-coding regions. In addition, as described below, other methods of downregulation of

gene expression which are well known in the art are contemplated and can be used with

the sequences of the present invention.

**Isolated Polynucleotides, Proteins, and Polypeptides**

15

          A first aspect of the present invention relates to isolated desaturase

polynucleotides. The polynucleotide sequences of the present invention include isolated

polynucleotides that are obtainable from genomic nucleic acid sequences.

          The invention provides a polynucleotide sequence identical over its entire length

20   to each sequence as set forth in the Sequence Listing. The polynucleotide includes non-

coding sequences, including for example, but not limited to, non-coding 5' and 3'

sequences, such as the transcribed, untranslated sequences, termination signals, ribosome

binding sites, sequences that stabilize mRNA, introns, polyadenylation signals, and

additional coding sequence that encodes additional amino acids. For example, a marker

25   sequence can be included to facilitate the purification of the fused polypeptide.

Polynucleotides of the present invention also include polynucleotides comprising a

structural gene and the naturally associated sequences that control gene expression.

The invention also includes polynucleotides of the formula:

$$X-(R_1)_n-(R_2)-(R_3)_n-Y$$
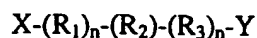
5   wherein, at the 5' end, X is hydrogen, and at the 3' end, Y is hydrogen or a metal, R_1 and

R_3 are any nucleic acid residue, n is an integer between 1 and 3000, preferably between 1

and 1000 and R_2 is a nucleic acid sequence of the invention, particularly a nucleic acid

sequence selected from the group set forth in the Sequence Listing and preferably SEQ ID

NOs: 1 - 8, and 23 - 29. In the formula, R_2 is oriented so that its 5' end residue is at the

10   left, bound to R_1, and its 3' end residue is at the right, bound to R_3. Any stretch of nucleic

acid residues denoted by either R group, where R is greater than 1, may be either a

heteropolymer or a homopolymer, preferably a heteropolymer.

Further preferred embodiments of the invention that are at least 50%, 60%, or 70%

identical over their entire length to a polynucleotide of the invention, and polynucleotides

15   that are complementary to such polynucleotides. More preferable are polynucleotides that

comprise a region that is at least 80% identical over its entire length to a polynucleotide of

the invention and polynucleotides that are complementary thereto. In this regard,

polynucleotides at least 90% identical over their entire length are particularly preferred,

those at least 95% identical are especially preferred. Further, those with at least 97%

20   identity are highly preferred and those with at least 98% and 99% identity are particularly

highly preferred, with those at least 99% being the most highly preferred.

Preferred embodiments are polynucleotides that are obtained from genomic

polynucleotide sequences and set forth in the Sequence Listing.

The invention further relates to polynucleotides that hybridize to the above-

25   described sequences. In particular, the invention relates to polynucleotides that hybridize

8

under stringent conditions to the above-described polynucleotides. As used herein, the

terms "stringent conditions" and "stringent hybridization conditions" mean that

hybridization will generally occur if there is at least 95% and preferably at least 97%

identity between the sequences. An example of stringent hybridization conditions is

5 overnight incubation at 42°C in a solution comprising 50% formamide, 5x SSC (150 mM

NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's

solution, 10% dextran sulfate, and 20 micrograms/milliliter denatured, sheared salmon

sperm DNA, followed by washing the hybridization support in 0.1x SSC at approximately

65°C. Other hybridization and wash conditions are well known and are exemplified in

10 Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, cold Spring

Harbor, NY (1989), particularly Chapter 11.

The invention also provides a polynucleotide consisting essentially of a

polynucleotide sequence obtainable by screening an appropriate library containing the

complete gene for a polynucleotide sequence set forth in the Sequence Listing under

15 stringent hybridization conditions with a probe having the sequence of said

polynucleotide sequence or a fragment thereof; and isolating said polynucleotide

sequence. Fragments useful for obtaining such a polynucleotide include, for example,

probes and primers as described herein.

As discussed herein regarding polynucleotide assays of the invention, for example,

20 polynucleotides of the invention can be used as a hybridization probe for RNA, cDNA, or

genomic DNA to isolate full length cDNAs or genomic clones encoding a polypeptide

and to isolate cDNA or genomic clones of other genes that have a high sequence

similarity to a polynucleotide set forth in the Sequence Listing. Such probes will

generally comprise at least 15 bases. Preferably such probes will have at least 30 bases

and can have at least 50 bases. Particularly preferred probes will have between 30 bases

and 50 bases, inclusive.

The region of each gene that comprises or is comprised by a polynucleotide

sequence set forth in the Sequence Listing may be isolated by screening using a DNA

5   sequence provided in the Sequence Listing to synthesize an oligonucleotide probe. A

labeled oligonucleotide having a sequence complementary to that of a polynucleotide of

the invention is then used to screen a library of cDNA, genomic DNA or mRNA to

identify members of the library which hybridize to the probe. For example, synthetic

oligonucleotides are prepared which correspond to the desaturase promoter and intron

10   sequences. In particular, screening of cDNA libraries in phage vectors is useful in such

methods due to lower levels of background hybridization.

Typically, a desaturase sequence obtainable from the use of nucleic acid probes

will show 60-70% sequence identity between the target desaturase sequence and the

encoding sequence used as a probe. However, lengthy sequences with as little as 50-60%

15   sequence identity may also be obtained. The nucleic acid probes may be a lengthy

fragment of the nucleic acid sequence, or may also be a shorter, oligonucleotide probe.

When longer nucleic acid fragments are employed as probes (greater than about 100 bp),

one may screen at lower stringencies in order to obtain sequences from the target sample

which have 20-50% deviation (i.e., 50-80% sequence homology) from the sequences used

20   as probe. Oligonucleotide probes can be considerably shorter than the entire nucleic acid

sequence encoding an desaturase enzyme, but should be at least about 10, preferably at

least about 15, and more preferably at least about 20 nucleotides. A higher degree of

sequence identity is desired when shorter regions are used as opposed to longer regions.

It may thus be desirable to identify regions of highly conserved amino acid sequence to

25   design oligonucleotide probes for detecting and recovering other related desaturase genes.

Shorter probes are often particularly useful for polymerase chain reactions (PCR),

especially when highly conserved sequences can be identified. *(See*, Gould, *et al., PNAS*

USA (1989) *86*:1934-1938.).

"Identity", as is well understood in the art, is a relationship between two or more

5   polypeptide sequences or two or more polynucleotide sequences, as determined by

comparing the sequences. In the art, "identity" also means the degree of sequence

relatedness between polypeptide or polynucleotide sequences, as determined by the match

between strings of such sequences. "Identity" can be readily calculated by known

methods including, but not limited to, those described in *Computational Molecular*

10   *Biology*, Lesk, A.M., ed., Oxford University Press, New York (1988); *Biocomputing:*

Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993;

Computer Analysis of Sequence Data, Part I, Griffin, A.M. and Griffin, H.G., eds.,

Humana Press, New Jersey (1994); *Sequence Analysis in Molecular Biology*, von Heinje,

G., Academic Press (1987); *Sequence Analysis Primer*, Gribskov, M. and Devereux, J.,

15   eds., Stockton Press, New York (1991); and Carillo, H., and Lipman, D., SIAM *J Applied*

Math, 48:1073 (1988). Methods to determine identity are designed to give the largest

match between the sequences tested. Moreover, methods to determine identity are

codified in publicly available programs. Computer programs which can be used to

determine identity between two sequences include, but are not limited to, GCG

20   (Devereux, J., et al., *Nucleic Acids Research* 12(1):387 (1984); suite of five BLAST

programs, three designed for nucleotide sequences queries (BLASTN, BLASTX, and

TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN)

(Coulson, *Trends in Biotechnology, 12:* 76-80 (1994); Birren, *et al., Genome Analysis, 1:*

543-559 (1997)). The BLAST X program is publicly available from NCBI and other

25   sources (*BLAST Manual*, Altschul, S., *et al.*, NCBI NLM NIH, Bethesda, MD 20894;

Altschul, S., *et al.*, *J. Mol. Biol.*, 215:403-410 (1990)). The well known Smith Waterman

algorithm can also be used to determine identity.

Parameters for polypeptide sequence comparison typically include the following:

Algorithm: Needleman and Wunsch, *J. Mol. Biol.* 48:443-453 (1970)

5   Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, *Proc. Natl.*

Acad. Sci USA 89:10915-10919 (1992)

Gap Penalty: 12

Gap Length Penalty: 4

A program which can be used with these parameters is publicly available as the

10   "gap" program from Genetics Computer Group, Madison Wisconsin. The above

parameters along with no penalty for end gap are the default parameters for peptide

comparisons.

Parameters for polynucleotide sequence comparison include the following:

Algorithm: Needleman and Wunsch, J. Mol. Biol. 48:443-453 (1970)

15   Comparison matrix: matches = +10; mismatches = 0

Gap Penalty: 50

Gap Length Penalty: 3

A program which can be used with these parameters is publicly available as the

"gap" program from Genetics Computer Group, Madison Wisconsin. The above

20   parameters are the default parameters for nucleic acid comparisons.

For immunological screening, antibodies to the protein can be prepared by

injecting rabbits or mice with the purified protein or portion thereof, such methods of

preparing antibodies being well known to those in the art. Either monoclonal or

polyclonal antibodies can be produced, although typically polyclonal antibodies are more

25   useful for gene isolation. Western analysis may be conducted to determine that a related

12

protein is present in a crude extract of the desired plant species, as determined by cross-reaction with the antibodies to the encoded proteins. When cross-reactivity is observed, genes encoding the related proteins are isolated by screening expression libraries representing the desired plant species. Expression libraries can be constructed in a variety

5   of commercially available vectors, including lambda gt11, as described in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual*, Second Edition (1989) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

Plant Constructs and Methods of Use

10

Of particular interest is the use of the polynucleotide sequences in recombinant DNA constructs to direct the transcription of the desaturase genomic sequences of the present invention in a host plant cell. The expression constructs generally comprise a promoter functional in a host plant cell operably linked to a nucleic acid sequence of the

15   present invention and a transcriptional termination region functional in a host plant cell.

Those skilled in the art will recognize that there are a number of promoters which are functional in plant cells, and have been described in the literature. Chloroplast and plastid specific promoters, chloroplast or plastid functional promoters, and chloroplast or plastid operable promoters are also envisioned.

20   One set of promoters are constitutive promoters such as the CaMV35S or FMV35S promoters that yield high levels of expression in most plant organs. Enhanced or duplicated versions of the CaMV35S and FMV35S promoters are useful in the practice of this invention (Odell, *et al.* (1985) *Nature* 313:810-812; Rogers, U.S. Patent Number 5,378, 619). In addition, it may also be preferred to bring about expression of the

25   sequences of the present invention in specific tissues of the plant, such as leaf, stem, root,

13

tuber, seed, fruit, etc., and the promoter chosen should have the desired tissue and

developmental specificity.

Of particular interest is the expression of the nucleic acid sequences of the present

invention from transcription initiation regions which are preferentially expressed in a

5   plant seed tissue. Examples of such seed preferential transcription initiation sequences

include those sequences derived from sequences encoding plant storage protein genes or

from genes involved in fatty acid biosynthesis in oilseeds. Examples of such promoters

include the 5' regulatory regions from such genes as napin (Kridle*t al., Seed Sci. Res.*

*1:*209:219 (1991)), phaseolin, zein, soybean trypsin inhibitor, ACP, stearoyl-ACP

10   desaturase, soybean α' subunit of β-conglycinin (soy 7s, (Chen *et al., Proc. Natl. Acad.*

Sci., 83:8560-8564 (1986))) and oleosin.

It may be advantageous to direct the localization of proteins conferring desaturase

to a particular subcellular compartment, for example, to the mitochondrion, endoplasmic

reticulum, vacuoles, chloroplast or other plastidic compartment. For example, where the

15   genes of interest of the present invention will be targeted to plastids, such as chloroplasts,

for expression, the constructs will also employ the use of sequences to direct the gene to

the plastid. Such sequences are referred to herein as chloroplast transit peptides (CTP) or

plastid transit peptides (PTP). In this manner, where the gene of interest is not directly

inserted into the plastid, the expression construct will additionally contain a gene

20   encoding a transit peptide to direct the gene of interest to the plastid. The chloroplast

transit peptides may be derived from the gene of interest, or may be derived from a

heterologous sequence having a CTP. Such transit peptides are known in the art. See, for

example, Von Heijne *et al.* (1991) *Plant Mol. Biol. Rep. 9:*104-126; Clark *et al.* (1989) *J.*

*Biol. Chem. 264:*17544-17550; della-Cioppa *et al.* (1987) *Plant Physiol. 84:*965-968;

Romer *et al.* (1993) *Biochem. Biophys. Res Commun. 196*:1414-1421; and, Shah *et al.*

(1986) *Science 233*:478-481.

Depending upon the intended use, the constructs may contain the entire genomic

nucleic acid sequence or a particular non-coding region of such a sequence or a portion of

5   such sequences. For example, where antisense inhibition of a given desaturase protein is

desired, the entire sequence is not required. Furthermore, where desaturase sequences

used in constructs are intended for use as probes, it may be advantageous to prepare

constructs containing only a particular portion of a desaturase sequence, for example a

sequence which encodes a highly conserved desaturase region.

10   The skilled artisan will recognize that there are various methods for the inhibition

of expression of endogenous sequences in a host cell. Such methods include, but are not

limited to, antisense suppression (Smith, *et al.* (1988) *Nature* 334:724-726) , co-

suppression (Napoli, *et al.* (1989) *Plant Cell* 2:279-289), ribozymes (PCT Publication

WO 97/10328), combinations of sense and antisense (Waterhouse, *et al.* (1998) *Proc.*

15   *Natl. Acad. Sci. USA* 95:13959-13964), promoter silencing (Park, *et al.* (1996) *Plant J.*

9(2):183-194), DNA binding proteins (Beerli, *et al.* (1997) *Proc. Natl. Acad. Sci. USA*,

95:14628-14633; and Liu, *et al.* (1998) *Proc. Natl. Acad. Sci. USA*, 94:5525-5530).

Methods for the suppression of endogenous sequences in a host cell typically employ the

transcription or transcription and translation of at least a portion of the sequence to be

20   suppressed. Such sequences may be homologous to coding as well as non-coding regions

of the endogenous sequence.

Regulatory transcript termination regions may be provided in plant expression

constructs of this invention as well. Transcript termination regions may be provided by

the DNA sequence encoding the desaturase or a convenient transcription termination

25   region derived from a different gene source, for example, the transcript termination region

15

which is naturally associated with the transcript initiation region. The skilled artisan will recognize that any convenient transcript termination region which is capable of terminating transcription in a plant cell may be employed in the constructs of the present invention.

5   Alternatively, constructs may be prepared to direct the expression of the desaturase sequences directly from the host plant cell plastid. Such constructs and methods are known in the art and are generally described, for example, in Svab, *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:8526-8530 and Svab and Maliga (1993) *Proc. Natl. Acad. Sci. USA* 90:913-917 and in U.S. Patent Number 5,693,507.

10   A plant cell, tissue, organ, or plant into which the recombinant DNA constructs containing the expression constructs have been introduced is considered transformed, transfected, or transgenic. A transgenic or transformed cell or plant also includes progeny of the cell or plant and progeny produced from a breeding program employing such a transgenic plant as a parent in a cross and exhibiting an altered phenotype resulting from

15   the presence of a desaturase nucleic acid sequence.

           Plant expression or transcription constructs having a desaturase polynucleotide of the present invention as the DNA sequence of interest for increased or decreased expression thereof may be employed with a wide variety of plant life, particularly, plant life involved in the production of vegetable oils for edible and industrial uses. Most

20   especially preferred are temperate oilseed crops. Plants of interest include, but are not limited to, rapeseed (Canola and High Erucic Acid varieties), sunflower, safflower, cotton, soybean, peanut, coconut and oil palms, and corn. Depending on the method for introducing the recombinant constructs into the host cell, other DNA sequences may be required. Importantly, this invention is applicable to dicotyledons and monocotyledons

species alike and will be readily applicable to new and/or improved transformation and

regulation techniques.

Of particular interest, is the use of plant desaturase promoter and/or intron

constructs in plants to produce plants or plant parts, including, but not limited to leaves,

5   stems, roots, reproductive, and seed, with a modified fatty acid composition. Of

particular interest in the desaturase promoter and/or intron constructs is the use of the

promoter and/or intron sequences of the Δ-12 and Δ-15 desaturase genomic sequences in

sense or antisense orientations for the modification of fatty acid compositions in host

cells.

10   The polynucleotides of the present invention can be used in the preparation of

constructs for use in a variety of host cells. Host for use in the present invention include,

but are not limited to plant cells, bacterial cells, fungal cells (including yeast), insect cells,

and mammalian cells.

For example, to confirm the activity and specificity of the proteins encoded by the

15   identified nucleic acid sequences as desaturase enzymes, *in vitro* assays can be performed

in insect cell cultures using baculovirus expression systems. Such baculovirus expression

systems are known in the art and are described by Lee, *et al.* U.S. Patent Number

5,348,886, the entirety of which is herein incorporated by reference.

The method of transformation in obtaining such transgenic plants is not critical to

20   the instant invention, and various methods of plant transformation are currently available.

Furthermore, as newer methods become available to transform crops, they may also be

directly applied hereunder. For example, many plant species naturally susceptible to

Agrobacterium infection may be successfully transformed via tripartite or binary vector

methods of *Agrobacterium* mediated transformation. In many instances, it will be

25   desirable to have the construct bordered on one or both sides by T-DNA, particularly

17

having the left and right borders, more particularly the right border. This is particularly

useful when the construct uses *A. tumefaciens* or *A. rhizogenes* as a mode for

transformation, although the T-DNA borders may find use with other modes of

transformation. In addition, techniques of microinjection, DNA particle bombardment,

5   and electroporation have been developed which allow for the transformation of various

monocot and dicot plant species.

Normally, included with the DNA construct will be a structural gene having the

necessary regulatory regions for expression in a host and providing for selection of

transformant cells. The gene may provide for resistance to a cytotoxic agent, e.g.

10   antibiotic, heavy metal, toxin, etc., complementation providing prototrophy to an

auxotrophic host, viral immunity or the like. Depending upon the host species the

expression construct or components thereof are introduced, one or more markers may be

employed, where different conditions for selection are used for the different hosts.

Where *Agrobacterium* is used for plant cell transformation, a vector may be used

15   which may be introduced into the *Agrobacterium* host for homologous recombination

with T-DNA or the Ti- or Ri-plasmid present in the *Agrobacterium* host. The Ti- or Ri-

plasmid containing the T-DNA for recombination may be armed (capable of causing gall

formation) or disarmed (incapable of causing gall formation), the latter being permissible,

so long as the *vir* genes are present in the transformed *Agrobacterium* host. The armed

20   plasmid can give a mixture of normal plant cells and gall.

In some instances where *Agrobacterium* is used as the vehicle for transforming

host plant cells, the expression or transcription construct bordered by the T-DNA border

region(s) will be inserted into a broad host range vector capable of replication in *E. coli*

and *Agrobacterium*, there being broad host range vectors described in the literature.

25   Commonly used is pRK2 or derivatives thereof. See, for example, Ditta, *et al.*, (*Proc.*

Nat. Acad. Sci., U.S.A. (1980) *77*:7347-7351) and EPA 0 120 515, which are incorporated

herein by reference. Alternatively, one may insert the sequences to be expressed in plant

cells into a vector containing separate replication sequences, one of which stabilizes the

vector in *E. coli*, and the other in *Agrobacterium*. See, for example, McBride and

5   Summerfelt (*Plant Mol. Biol.* (1990) *14*:269-276), wherein the pRiHRI (Jouanin, *et al.,*

Mol. Gen. Genet. (1985) 201:370-374) origin of replication is utilized and provides for

added stability of the plant expression vectors in host *Agrobacterium* cells.

Included with the expression construct and the T-DNA will be one or more

markers, which allow for selection of transformed Agrobacterium and transformed plant

10   cells. A

number of markers have been developed for use with plant cells, such as resistance to

chloramphenicol, kanamycin, the aminoglycoside G418, hygromycin, or the like. The

particular marker employed is not essential to this invention, one or another marker being

preferred depending on the particular host and the manner of construction.

15   For transformation of plant cells using *Agrobacterium*, explants may be combined

and incubated with the transformed *Agrobacterium* for sufficient time for transformation,

the bacteria killed, and the plant cells cultured in an appropriate selective medium. Once

callus forms, shoot formation can be encouraged by employing the appropriate plant

hormones in accordance with known methods and the shoots transferred to rooting

20   medium for regeneration of plants. The plants may then be grown to seed and the seed

used to establish repetitive generations and for isolation of vegetable oils.

For the alteration of unsaturated fatty acid production in a host cell, a second

expression construct can be used in accordance with the present invention. For example,

the desaturase expression construct can be introduced into a host cell in conjunction with

a second expression construct having a nucleotide sequence for a protein involved in

fatty acid biosynthesis.

There are several possible ways to obtain the plant cells of this invention which

contain multiple expression constructs. Any means for producing a plant comprising a

5   construct having a DNA sequence encoding the expression construct of the present

invention, and at least one other construct having another DNA sequence encoding an

enzyme are encompassed by the present invention. For example, the expression construct

can be used to transform a plant at the same time as the second construct either by

inclusion of both expression constructs in a single transformation vector or by using

10   separate vectors, each of which express desired genes. The second construct can be

introduced into a plant which has already been transformed with the desaturase expression

construct, or alternatively, transformed plants, one expressing the desaturase construct and

one expressing the second construct, can be crossed to bring the constructs together in the

same plant.

15   The invention now being generally described, it will be more readily understood

by reference to the following examples which are included for purposes of illustration

only and are not intended to limit the present invention.

EXAMPLES

20

Example 1 Cloning of Desaturase Genomic Sequences

1A. Soybean Δ12 Desaturase (fad2-1)

The soybean fad 2-1A sequence was identified by screening a soybean genomic

25   library using a soybean fad2-1 cDNA probe. Three putative soy fad 2-1 clones were

identified and plaque purified. Two of the three soy fad 2-1 clones were ligated into

pBluescript II KS+ (Stratagene) and sequenced. Both clones (14-1 and 11-12) were the

same and matched the soy fad 2-1 cDNA exactly. The sequence of the entire fad2-1A

clone is provided in SEQ ID NO: 1.

5   Prior to obtaining this full length clone, a portion of the fad2-1A genomic clone

was PCR amplified using PCR primers designed from the 5' untranslated sequence

(Primer 12506, 5'-ATACAA GCCACTAGGCAT-3', SEQ ID NO:9) and within the

cDNA (Primer 11698: 5'-GATTGGCCATGCAATGAGGGAAAAGG-3', SEQ ID

NO:10. The resulting PCR product, which contained the fad2-1A intron, was cloned into

10   the vector pCR 2.1 (Invitrogen) and sequenced. The soy fad 2-1A partial genomic clone

(SEQ ID NO:27) and its intron region (SEQ ID NO:2) were identified by comparison to

the soybean cDNA sequence using the Pustell comparison program in Macvector. The

intron sequence begins after the ATG start codon, and is 420 bases long.

          A second fad2-1 gene family member was also identified and cloned, and is

15   referred to herein as fad2-1B. The soy fad 2-1B partial genomic clone (SEQ ID NO:23)

(contains the promoter (base pairs 1 - 1704); 5'UTR (base pairs 1705 - 1782); intron#1

(base pairs 1786 - 2190); and a portion of the fad2-1B coding region (base pairs 1783-

1785 and 2191 - 2463)) and its intron region (SEQ ID NO:24) were identified by

comparison to the soybean cDNA sequence using the Pustell comparison program in

20   Macvector. The intron sequence begins after the ATG start codon and is 405 bases long.

1B. Soybean Δ15 Desaturase (fad3)

          The partial soybean fad 3 genomic sequence was PCR amplified from soybean

DNA using primers 10632, 5'-

25   CUACUACUACUACTCGAGACAAAGCCTTTAGCCTATG-3' (SEQ ID NO:11), and

10633: 5'-CAUCAUCAUCAUGGATCCCATGTC TCTCTATGCAAG-3' (SEQ ID

NO:12). The Expand Long Template PCR system (Boehringer Mannheim) was used

according to the manufacturers directions. The resulting PCR products were cloned into

the vector pCR 2.1 (Invitrogen) and sequenced. The soy fad 3 partial genomic clone

5   sequence and the intron regions were confirmed by comparisons to the soybean fad 3

cDNA sequence using the Pustell program in Macvector. From the identified partial

genomic soybean fad3 sequence (SEQ ID NO:3), seven introns were identified (SEQ ID

NO:4 (intron #1), SEQ ID NO:5 (intron #2), SEQ ID NO:6 (intron #3A), SEQ ID NO:7

(intron #4), SEQ ID NO:8 (intron #5), SEQ ID NO:25 (intron #3B) and SEQ ID NO:26

10   (intron #3C)). Intron #1 is 192 base pairs long and is located between positions 294 and

485, intron #2 is 348 base pairs long and is located between positions 576 and 923, intron

#3A is 142 base pairs long and is located between positions 991 and 1132, intron #3B is

98 base pairs long and is located between positions 1225 and 1322 , intron #3C is 115

base pairs long and is located between positions 1509 and 1623, intron #4 is 1231 base

15   pairs long and is located between positions 1705 and 2935, and intron #5 is 626 base pairs

long and is located between positions 3074 and 3699.

Example 2 Expression constructs

20   The soybean fad2-1A intron sequence was amplified via PCR using the fad2-1A

partial genomic clone (SEQ ID NO:27) as a template and primers 12701 (5'-

ACGAATTCCTCGAGGTAAA TTAAATTGTGCCTGC-3' (SEQ ID NO:13)) and

12702 (5'-GCGAGATCTATCG ATCTGTGTCAAAGTATAAAC-3' (SEQ ID NO:14)).

The resulting amplification products were cloned into the vector pCR 2.1 (Invitrogen) and

25   sequenced. The soyfad2-1A intron was then cloned into the expression cassette,

22

pCGN3892, in sense and antisense orientations. The vector pCGN3892 contains the

soybean 7S promoter and a pea RBCS 3'. Both gene fusions were then separately

ligated into pCGN9372, a vector that contains the CP4 gene regulated by the FMV

promoter. The resulting expression constructs (PCGN5469 sense and pCGN5471

5   antisense) were used for transformation of soybean using biolistic methods described

below.

The soybean fad2-1B intron sequence was amplified via PCR using the fad2-1B partial

genomic clone (SEQ ID NO:23) as a template and primers 13883 (5'-

10   GCGATCGATGTATGATGCTAAATTAAATTGTGCCTG -3' (SEQ ID NO:30)) and

13876 (5'- GCGGAATTCCTGTGTCAAAGTATAAAGAAG -3' (SEQ ID NO:31)). The

resulting amplification products were cloned into the vector pCR 2.1 (Invitrogen) and

sequenced. The soyfad2-1B intron was fused to the 3' end of the soy fad 2-1A intron in

plasmids pCGN5468 (contains the soybean 7S promoter fused to the soy fad2-1A intron

15   (sense) and a pea RBCS 3') or pCGN5470 (contains the soybean 7S promoter fused to

the soy fad2-1A intron (antisense) and a pea RBCS 3') in sense or antisense orientation

respectively. The resulting intron combo fusions were then ligated separately into

pCGN9372, a vector that contains the CP4 gene regulated by the FMV promoter. The

resulting expression constructs (pCGN5485, fad2-1A&B intron sense and pCGN5486,

20   fad2-1A&B intron antisense) were used for transformation of soybean using biolistic

methods described below.

Four of the seven introns identified from the soybean fad 3 genomic clone were

PCR amplified using the soy fad 3 partial genomic clone as template and primers as

25   follows: Intron #1, primers 12568: GATCGATGCCCGGGGTAATAATTTTTGTGT

(SEQ ID NO:15) and 12569: CACGCCTCGAGTGTTCAATTCAATCAATG (SEQ ID

NO:16); Intron #2, primers 12514: 5'-CACTCGAGTTAGTTCATACTGGCT (SEQ ID

NO:17) and 12515: 5'-CGCATCGATTGCAAAATCCATCAAA (SEQ ID NO:18);

Intron #4, primers 10926: 5'-

5   CUACUACUACUACTCGAGCGTAAATAGTGGGTGAACAC (SEQ ID NO:19) and

10927: 5'-CAUCAUCAUCAUCTCGAGGAATTCGTCCATTTTAGTACACC (SEQ ID

NO:20) ; Intron #5, primers 10928: 5'-CUACUACUACUACTCGAGGCGCGT

ACATTTTATTGCTTA (SEQ ID NO:21) and 10929: 5'-CAUCAUCAUCAUCT

CGAGGAATTCTGCAGTGAATCCAAATG (SEQ ID NO:22). The resulting PCR

10   products for each intron were cloned into the vector pCR 2.1 (Invitrogen) and sequenced.

Introns #1, #2, #4 and #5 were all ligated separately into the, pCGN3892, in sense or

antisense orientations. pCGN3892 contains the soybean 7S promoter and a pea RBCS

3'. These fusions were ligated into pCGN9372, a vector that contains the CP4 gene

regulated by the FMV promoter for transformation into soybean. The resulting expression

15   constructs (pCGN5455, fad3 intron#4 intron sense; pCGN5459, fad3 intron#4 intron

antisense; pCGN5456, fad3 intron#5 intron sense; pCGN5460, fad3 intron#5 intron

antisense; pCGN5466, fad3 intron#2 intron antisense; pCGN5473, fad3 intron#1 intron

antisense;) were used for transformation of soybean using biolistic methods described

below.

20   The soy fad3 Intron #3C and #4 were also PCR amplified from a second fad3 gene

family member, herein referred to as fad3-1B. The soy fad3-1B introns #3C and #4 were

PCR amplified from soybean DNA using the following primers, 5'

CATGCTTTCTGTGCTTCTC 3' (SEQ ID NO:32) and , 5'

GTTGATCCAACCATAGTCG 3' (SEQ ID NO:33). The PCR products were cloned into

the vector pCR 2.1 (Invitrogen) and sequenced. The sequences for the soy fad3-1B

introns #3C and #4 are provided in SEQ ID NOs:28 and 29.

Example 3 Plant Transformation and Analysis

5   Linear DNA fragments containing the expression constructs for sense and

antisense expression of the Δ12 and Δ15 desaturase introns were stably introduced into

soybean (Asgrow variety A4922) by the method of McCabe, et.al. (1988) Bio/Technology

6:923-926. Transformed soybean plants were identified by selection on media containing

glyphosate.

10   Fatty acid compositions were analyzed from seed of soybean lines

transformed with the intron expression constructs using gas chromatography. T2 pooled

seed and T2 single seed oil compositions demonstrate that the mono and polyunsaturated

fatty acid compositions were altered in the oil of seeds from transgenic soybean lines as

compared to that of the seed from non-transformed soybean. Table I provides a summary

15   of results which were obtained using the described constructs. These data clearly show

that sense and antisense expression of the non-coding regions of the desaturase gene

results in the modification of the fatty acid compositions. The data also shows that

introns can be used to obtain a variety of lines with varying fatty acid compositions.

Selections can be made from such lines depending on the desired relative fatty acid

20   composition. In addition, since each of the introns is able to modify the levels of each

fatty acid to varying extents, it is contemplated that combinations of introns can be used

depending on the desired compositions.

TABLE I

			Oleic	Linoleic	Linolenic
Fad 2	orientation	event			
wildtype (control)		5469-5 null T2 pool	18.15%	55.59%	7.97%
		10 seed average	13.89%	55.89%	9.067%
		5469-27 null T2 pool	19.15%	54.62%	9.32%
		A4922	15.75%	56.1%	8.75%
		5471-13 null T2 pool	17.02%	56.49%	9.08%
		10 seed average	13.86%	56.14%	9.49%
		A4922	14.95%	55.95%	9.07%
full length cDNA (control)	sense	5462-133 T2 pool	84%	2.17%	1.55%
		best 5462-133 T2 seed	84%	0.59%	1.76%
intron 1	sense	5469-6 T2 pool	29.93%	46.53%	
		5469-8 T2 pool	36.5%	42.11%	5.98%
		best 5469-6 T2 seed	44.41%	29.34%	6.68%
		best 5469-8 T2 seed	41.26%	33.16%	5.74%
		5469-14 T2 pool	61.06%	16.42%	7.75%
		5469-20 T2 pool	48.89%	31.61%	4.89%
		5469-22 T2 pool	80%	2.97%	4.78%
		best 5469-14 T2 seed	62.21%	11.97%	8.81%
		5485-3 T2 pool	63.54%	14.09%	7.32%
		5485-53 T2 pool	47.58%	27.64%	7.81%
	antisense	5471-8 T2 pool	31.05%	43.62%	7.07%
		5471-2 T2 pool	27.98%	48.88%	6.83%
		5471-26 T2 pool	32.66%	44.54%	6.76%
		best 5471-8 T2 seed	57.4%	23.37%	5.73%
		best 5471-2 T2 seed	28.08%	46.14%	6.52%
		best 5471-26 T2 seed	43.3%	34.15%	5.6%
		5486-33 T2 pool	32.37%	43.66%	6.87%
		5486-12 T2 pool	27.32%	46.97%	6.4%
		5486-40 T2 pool	26.79%	48.72%	6.55%
Fad 3					
wildtype (control)		5473-7 null T2 pool	15.65%	56.74%	9.55%
		A4922 T2 pool	19.84%	56.79%	7.48%
full length cDNA (control)	sense	5464-50 T2 pool	18.06%	62.03%	2.75%
		best 5464-50 T2 seed	17.08%	62.44%	1.72%
intron 1	antisense	5473-8 T2 pool	33.47%	45.97%	5.54%
		5473-1 T2 pool	33.34%	42.67%	7.59%
intron 2	antisense	5466-20 T2 pool	28.43%	48.83%	6.37%
		5466-16 T2 pool	27.61%	49.92%	5.96%
intron 4	sense	5455-19 T2 pool	40.35%	39.97%	4.61%
		5455-10 T2 pool	35.14%	43.59%	5.53%
		5455-57 T2 pool	38.04%	42.44%	5.24%
		5455-76 T2 pool	37.24%	42.42%	5.37%
		5455-107 T2 pool	36.44%	42.72%	5.62%
		best 5455-57 T2 seed	45.36%	35.55%	4.92%
		best 5455-76 T2 seed	35.3%	43.54%	5.53%
		best 5455-107 T2 seed	45.56%	34.85%	5.12%
	antisense	5459-2 T2 pool	34.5%	43.87%	5.59%
		5459-6 T2 pool	33.78%	44.12%	5.62%
		5459-20 T2 pool	28.26%	49.48%	5.5%
		best 5459-2 T2 seed	61.45%	23.45%	3.38%
		best 5459-6 T2 seed	53.51%	29.68%	3.53%

26

		best 5459-20 T2 seed	30%	50.55%	4.15%
intron 5	sense	5456-38 T2 pool	28.23%	49.59%	6.74%
		5456-62 T2 pool	28.94%	48.66%	6.25%
		best 5456-62 T2 seed	29.5%	43.69%	5.4%
	antisense	5460-9 T2 pool	29.78%	48.57%	5.54%
		5460-21 T2 pool	28.37%	49.79%	5.54%
		best 5460-21 T2 seed	35.18%	40.52%	5.33%

All publications and patent applications mentioned in this specification are

indicative of the level of skill of those skilled in the art to which this invention pertains.

5   All publications and patent applications are herein incorporated by reference to the same

extent as if each individual publication or patent application was specifically and

individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of

illustration and example for purposes of clarity of understanding, it will be obvious that

10   certain changes and modifications may be practiced within the scope of the appended

claims.

27

Claims

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO: 1;

5   b) a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 1

over the entire length of SEQ ID NO: 1;

c) a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 1

over the entire length of SEQ ID NO: 1;

d) a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 1

10   over the entire length of SEQ ID NO: 1;

e) a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 1

over the entire length of SEQ ID NO: 1;

f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 1 or a fragment thereof; and

15   g) a polynucleotide sequence complementary to the polynucleotide sequence of (a),

(b), (c), (d), (e), or (f).

2. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO: 2;

20   b) a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 2

over the entire length of SEQ ID NO: 2;

c) a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 2

over the entire length of SEQ ID NO: 2;

d) a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 2

25   over the entire length of SEQ ID NO: 2;

e) a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 2

over the entire length of SEQ ID NO: 2;

f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 2 or a fragment thereof; and

5   g) a polynucleotide sequence complementary to the polynucleotide sequence of (a),

(b), (c), (d), (e), or (f).

3.   An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO: 3;

10   b) a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 3

over the entire length of SEQ ID NO: 3;

c) a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 3

over the entire length of SEQ ID NO: 3;

d) a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 3

15   over the entire length of SEQ ID NO: 3;

e) a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 3

over the entire length of SEQ ID NO: 3;

f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 3 or a fragment thereof; and

20   g) a polynucleotide sequence complementary to the polynucleotide sequence of (a),

(b), (c), (d), (e), or (f).

4.   An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO: 4;

b) a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 4 over the entire length of SEQ ID NO: 4;

c) a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 4 over the entire length of SEQ ID NO: 4;

5 d) a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 4 over the entire length of SEQ ID NO: 4;

e) a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 4 over the entire length of SEQ ID NO: 4;

f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

10 NO: 4 or a fragment thereof; and

g) a polynucleotide sequence complementary to the polynucleotide sequence of (a), (b), (c), (d), (e), or (f).

5. An isolated polynucleotide selected from the group consisting of:

15 a) a polynucleotide comprising SEQ ID NO: 5;

b) a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 5 over the entire length of SEQ ID NO: 5;

c) a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 5 over the entire length of SEQ ID NO: 5;

20 d) a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 5 over the entire length of SEQ ID NO: 5;

e) a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 5 over the entire length of SEQ ID NO: 5;

f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

25 NO: 5 or a fragment thereof; and

g) a polynucleotide sequence complementary to the polynucleotide sequence of (a),

(b), (c), (d), (e), or (f).

6.   An isolated polynucleotide selected from the group consisting of:

5   a) a polynucleotide comprising SEQ ID NO: 6;

b) a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 6

over the entire length of SEQ ID NO: 6;

c) a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 6

over the entire length of SEQ ID NO: 6;

10   d) a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 6

over the entire length of SEQ ID NO: 6;

e) a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 6

over the entire length of SEQ ID NO: 6;

f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

15   NO: 6 or a fragment thereof; and

g) a polynucleotide sequence complementary to the polynucleotide sequence of (a),

(b), (c), (d), (e), or (f).

7.   An isolated polynucleotide selected from the group consisting of:

20   a) a polynucleotide comprising SEQ ID NO: 7;

b) a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 7

over the entire length of SEQ ID NO: 7;

c) a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 7

over the entire length of SEQ ID NO: 7;

d) a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 7

over the entire length of SEQ ID NO: 7;

e) a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 7

over the entire length of SEQ ID NO: 7;

5   f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 7 or a fragment thereof; and

g) a polynucleotide sequence complementary to the polynucleotide sequence of (a),

(b), (c), (d), (e), or (f).

10   8.   An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO: 8;

b) a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 8

over the entire length of SEQ ID NO: 8;

c) a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 8

15   over the entire length of SEQ ID NO: 8;

d) a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 8

over the entire length of SEQ ID NO: 8;

e) a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 8

over the entire length of SEQ ID NO: 8;

20   f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 8 or a fragment thereof; and

g) a polynucleotide sequence complementary to the polynucleotide sequence of (a),

(b), (c), (d), (e), or (f).

9.   An isolated polynucleotide selected from the group consisting of:

a)   a polynucleotide comprising SEQ ID NO: 23;

b)   a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 23

over the entire length of SEQ ID NO: 23;

5   c)   a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 23

over the entire length of SEQ ID NO: 23;

d)   a polynucleotide sequence having at least 90% identity to that of SEQ ID NO:23

over the entire length of SEQ ID NO: 23;

e)   a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 23

10   over the entire length of SEQ ID NO: 23;

f)   an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 23 or a fragment thereof; and

g)   a polynucleotide sequence complementary to the polynucleotide sequence of (a),

(b), (c), (d), (e), or (f).

15

10.   An isolated polynucleotide selected from the group consisting of:

a)   a polynucleotide comprising SEQ ID NO: 24;

b)   a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 24

over the entire length of SEQ ID NO: 24;

20   c)   a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 24

over the entire length of SEQ ID NO: 24;

d)   a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 24

over the entire length of SEQ ID NO: 24;

e)   a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 24

25   over the entire length of SEQ ID NO: 24;

f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 24 or a fragment thereof; and

g) a polynucleotide sequence complementary to the polynucleotide sequence of (a),

(b), (c), (d), (e), or (f).

5

11.   An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO: 25;

b) a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 25

over the entire length of SEQ ID NO: 25;

10   c) a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 25

over the entire length of SEQ ID NO: 25;

d) a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 25

over the entire length of SEQ ID NO: 25;

e) a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 25

15   over the entire length of SEQ ID NO: 25;

f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 25 or a fragment thereof; and

a polynucleotide sequence complementary to the polynucleotide sequence of (a),

(b), (c), (d), (e), or (f).

20

12.   An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO: 26;

b) a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 26

over the entire length of SEQ ID NO: 26;

34

c) a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 26

over the entire length of SEQ ID NO: 26;

d) a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 26

over the entire length of SEQ ID NO: 26;

5   e) a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 26

over the entire length of SEQ ID NO: 26;

f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 26 or a fragment thereof; and

a polynucleotide sequence complementary to the polynucleotide sequence of (a),

10   (b), (c), (d), (e), or (f).

13.   An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO: 27;

b) a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 27

15   over the entire length of SEQ ID NO: 27;

c) a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 27

over the entire length of SEQ ID NO: 27;

d) a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 27

over the entire length of SEQ ID NO: 27;

20   e) a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 27

over the entire length of SEQ ID NO: 27;

f) an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 27 or a fragment thereof; and

a polynucleotide sequence complementary to the polynucleotide sequence of (a),

25   (b), (c), (d), (e), or (f).

35

14.   An isolated polynucleotide selected from the group consisting of:

a)   a polynucleotide comprising SEQ ID NO: 28;

b)   a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 28

5   over the entire length of SEQ ID NO: 28;

c)   a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 28

over the entire length of SEQ ID NO: 28;

d)   a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 28

over the entire length of SEQ ID NO: 28;

10   e)   a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 28

over the entire length of SEQ ID NO: 28;

f)   an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 28 or a fragment thereof; and

a polynucleotide sequence complementary to the polynucleotide sequence of (a),

15   (b), (c), (d), (e), or (f).

15.   An isolated polynucleotide selected from the group consisting of:

a)   a polynucleotide comprising SEQ ID NO: 29;

b)   a polynucleotide sequence having at least 70% identity to that of SEQ ID NO: 29

20   over the entire length of SEQ ID NO: 29;

c)   a polynucleotide sequence having at least 80% identity to that of SEQ ID NO: 29

over the entire length of SEQ ID NO: 29;

d)   a polynucleotide sequence having at least 90% identity to that of SEQ ID NO: 29

over the entire length of SEQ ID NO: 29;

e)   a polynucleotide sequence having at least 95% identity to that of SEQ ID NO: 29

over the entire length of SEQ ID NO: 29;

f)   an polynucleotide sequence that hybridizes, under stringent conditions, to SEQ ID

NO: 29 or a fragment thereof; and

a polynucleotide sequence complementary to the polynucleotide sequence of (a),

(b), (c), (d), (e), or (f).

16.   An intron obtained from a genomic polynucleotide sequence selected from

the group consisting of:

a)   a genomic polynucleotide sequence having at least 70% identity to coding

regions of SEQ ID NO: 1 over the entire coding regions of SEQ ID NO:1;

b)   a genomic polynucleotide sequence having at least 80% identity to coding

regions of SEQ ID NO: 1 over the entire coding regions of SEQ ID NO:1;

c)   a genomic polynucleotide sequence having at least 90% identity to coding

regions of SEQ ID NO: 1 over the entire coding regions of SEQ ID NO:1; and

d)   a genomic polynucleotide sequence having at least 95% identity to coding

regions of SEQ ID NO: 1 over the entire coding regions of SEQ ID NO:1.

17.   An intron obtained from a genomic polynucleotide sequence selected from

the group consisting of:

a)   a genomic polynucleotide sequence having at least 70% identity to coding

regions of SEQ ID NO: 3 over the entire coding regions of SEQ ID NO:3;

b)   a genomic polynucleotide sequence having at least 80% identity to coding

regions of SEQ ID NO: 3 over the entire coding regions of SEQ ID NO:3;

c)   a genomic polynucleotide sequence having at least 90% identity to coding

regions of SEQ ID NO: 3 over the entire coding regions of SEQ ID NO:3; and

d)   a genomic polynucleotide sequence having at least 95% identity to coding

regions of SEQ ID NO: 3 over the entire coding regions of SEQ ID NO:3.

5

18.   An intron obtained from a genomic polynucleotide sequence selected from

the group consisting of:

a)   a genomic polynucleotide sequence having at least 70% identity to coding

regions of SEQ ID NO:23 over the entire coding regions of SEQ ID NO:23;

10   b)   a genomic polynucleotide sequence having at least 80% identity to coding

regions of SEQ ID NO:23 over the entire coding regions of SEQ ID NO:23;

c)   a genomic polynucleotide sequence having at least 90% identity to coding

regions of SEQ ID NO:23 over the entire coding regions of SEQ ID NO:23; and

d)   a genomic polynucleotide sequence having at least 95% identity to coding

15   regions of SEQ ID NO:23 over the entire coding regions of SEQ ID NO:23.

19.   An intron obtained from a genomic polynucleotide sequence selected from

the group consisting of:

a)   a genomic polynucleotide sequence having at least 70% identity to coding

20   regions of SEQ ID NO:27 over the entire coding regions of SEQ ID NO:27;

b)   a genomic polynucleotide sequence having at least 80% identity to coding

regions of SEQ ID NO:27 over the entire coding regions of SEQ ID NO:27;

c)   a genomic polynucleotide sequence having at least 90% identity to coding

regions of SEQ ID NO:27 over the entire coding regions of SEQ ID NO:27; and

d)   a genomic polynucleotide sequence having at least 95% identity to coding

regions of SEQ ID NO:27 over the entire coding regions of SEQ ID NO:27.

20.   A recombinant DNA construct comprising at least one of the

5   polynucleotide sequences of Claims 1-19.

21.   The recombinant DNA construct according to Claim 20, wherein said

polynucleotide sequence is an intron sequence.

10   22.   The recombinant DNA construct according to Claim 21, wherein said

intron sequence is in an orientation selected from the group consisting of sense and

antisense.

23.   A plant cell comprising the DNA construct of 20.

15

24.   A plant comprising the cell of Claim 24.

25.   A method of modifying the fatty acid composition in a plant cell, said

method comprising:

20   transforming a plant cell with the construct of Claim 21 and,

growing said cell under conditions wherein transcription of said polynucleotide

sequence is initiated,

whereby said fatty acid composition of said cell is modified.

26.   The method according to Claim 25, wherein said modification is selected

from the group consisting of an increase in oleic acid, a decrease in linolenic acid and a

decrease in linoleic acid.

5   27.   The method according to Claim 25, wherein said polynucleotide sequence

is an intron sequence.

28.   The method according to Claim 27, wherein said intron sequence is in an

orientation selected from the group consisting of sense and antisense.

10

29.   A plant cell produced by the method of claim 25.

30.   A method of inhibiting gene expression in a plant cell, comprising the steps of:

transforming a plant cell with a DNA construct comprising a non-coding region of a gene to

15   be inhibited positioned in a sense orientation; and

growing said cell under conditions wherein transcription of said non-coding region is

initiated, whereby said expression of said gene is inhibited.

31.   The method according to claim 30 wherein said non-coding region is selected from

20   the group consisting of an intron, a promoter region, a 5' untranslated region and portions of such

sequences.

32.   The method according to claim 30 wherein said non-coding region is selected from

the group consisting of SEQ ID NO:2, SEQ ID NO:24, SEQ ID NO:4, SEQ ID NO:5, SEQ ID

NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:28 and SEQ ID

NO:29.

33.   The method according to claim 30 wherein said gene encodes desaturase.

5

34.   A method of inhibiting desaturase expression in a plant cell, comprising the steps of:

transforming a plant cell with a DNA construct comprising a non-coding region of a gene

encoding desaturase positioned in an antisense orientation; and

growing said cell under conditions wherein transcription of said non-coding region is

10   initiated, whereby said expression of said desaturase is inhibited.

35.   The method according to claim 34 wherein said non-coding region is selected from

the group consisting of an intron, a promoter region, a 5' untranslated region and portions of such

sequences.

15

36.   The method according to claim 34 wherein said non-coding region is selected from

the group consisting of SEQ ID NO:2, SEQ ID NO:24, SEQ ID NO:4, SEQ ID NO:5, SEQ ID

NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:28 and SEQ ID

NO:29.

20

37.   A plant cell produced according to the method of claim 34.

38.   The plant cell according to claim 37 wherein said cell includes an oil composition

comprising at least about 80 - 85% oleic acid, no more than about 1 - 2% linoleic acid and no more

25   than about 1-3% linolenic acid.

39.   The plant cell according to claim 37 wherein said cell includes an oil composition

comprising at least about 50 - 75% oleic acid, no more than about 10 - 30% linoleic acid and no

more than about 1-3% linolenic acid.

5

40.   A method of modifying fatty acid composition in a plant cell, comprising the steps

of:

transforming a plant cell with a DNA construct comprising a non-coding region of a gene

encoding desaturase positioned in a sense or an antisense orientation; and

10   growing said cell under conditions wherein transcription of said non-coding region is

initiated, whereby said expression of said desaturase is inhibited.

41.   The method according to claim 40 wherein said non-coding region is selected from

the group consisting of an intron, a promoter region, a 5' untranslated region and portions of such

15   sequences.

42.   The method according to claim 41 wherein said non-coding region is selected from

the group consisting of SEQ ID NO:2, SEQ ID NO:24, SEQ ID NO:4, SEQ ID NO:5, SEQ ID

NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:28 and SEQ ID

20   NO:29.

43.   A plant cell produced according to the method of claim 40.

44.   The plant cell according to claim 43 wherein said cell includes an oil composition comprising at least about 80 - 85% oleic acid, no more than about 1 - 2% linoleic acid and no more than about 1-3% linolenic acid.

5   45.   The plant cell according to claim 43 wherein said cell includes an oil composition comprising at least about 50 - 75% oleic acid, no more than about 10 - 30% linoleic acid and no more than about 1-3% linolenic acid.

46.   A method of inhibiting desaturase expression in a plant cell, comprising the steps of:

10   transforming a plant cell with a DNA construct comprising a nucleic acid sequence capable of binding to or cleaving a non-coding region of a gene encoding desaturase; and

growing said cell under conditions wherein transcription of said nucleic acid sequence is initiated, whereby said expression of said desaturase is inhibited.

15   47.   A plant cell produced according to the method of claim 46.

48.   A oil seed composition comprising:

at least about 50 - 75% oleic acid, no more than about 10 - 30% linoleic acid and no more than about 1-3% linolenic acid.

20

SEQUENCE LISTING

<110> Fillatti, Joanne J.

<120> Nucleic Acid Sequences and Methods of
     Use for the Production of Plants with Modified
     Polyunsaturated Fatty Acids

<130> 17138/00/WO

<140> new application
<141> 2000-08-11

<150> 60/151,224
<151> 1999-08-26

<150> 60/172,128
<151> 1999-12-17

<160> 33

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 4497
<212> DNA
<213> Glycine max

<400> 1

```
cttgcttggt aacaacgtcg tcaagttatt attttgttct tttttttttt atcatatttc      60
ttattttgtt ccaagtatgt catattttga tccatcttga caagtagatt gtcatgtagg     120
aataggaata tcactttaaa ttttaaagca ttgattagtc tgtaggcaat attgtcttct     180
tcttcctcct tattaatatt ttttattctg ccttcaatca ccagttatgg gagatggatg     240
taatactaaa taccatagtt gttctgcttg aagtttagtt gtatagttgt tctgcttgaa     300
gtttagttgt gtgtaatgtt tcagcgttgg cttcccctgt aactgctaca atggtactga     360
atatatattt tttgcattgt tcatttttt cttttactta atcttcattg ctttgaaatt     420
aataaaacaa aaagaaggac cgaatagttt gaagtttgaa ctattgccta ttcatgtaac     480
ttattcaccc aatcttatat agttttctg gtagagatca ttttaaattg aaggatataa     540
attaagagga aatacttgta tgtgatgtgt ggcaatttgg aagatcatgc gtagagagtt     600
taatggcagg ttttgcaaat tgacctgtag tcataattac actgggccct ctcggagttt     660
tgtgccttt tgttgtcgct gtgtttggtt ctgcatgtta gcctcacaca gatatttagt     720
agttgttgtt ctgcatataa gcctcacacg tatactaaac gagtgaacct caaaatcatg     780
gccttacacc tattgagtga aattaatgaa cagtgcatgt gagtatgtga ctgtgacaca     840
accccgggtt ttcatattgc aatgtgctac tgtggtgatt aaccttgcta cactgtcgtc     900
cttgtttgtt tccttatgta tattgatacc ataaattatt actagtatat cattttatat     960
tgtccatacc attacgtgtt tatagtctct ttatgacatg taattgaatt ttttaattat    1020
aaaaaataat aaaacttaat tacgtactat aaagagatgc tcttgactag aattgtgatc    1080
tcctagtttc ctaaccatat actaatattt gcttgtattg atagcccctc cgttcccaag    1140
agtataaaac tgcatcgaat aatacaagcc actaggcatg gtaaattaaa ttgtgcctgc    1200
acctcgggat atttcatgtg gggttcatca tatttgttga ggaaaagaaa ctcccgaaat    1260
tgaattatgc atttatatat cctttttcat ttctagattt cctgaaggct taggtgtagg    1320
cacctagcta gtagctacaa tatcagcact tctctctatt gataaacaat tggctgtaat    1380
gccgcagtag aggacgatca caacatttcg tgctggttac ttttttgtttt atggtcatga    1440
tttcactctc tctaatctct ccattcattt tgtagttgtc attatcttta gatttttcac    1500
```

- 1 -

```
tacctggttt aaaattgagg gattgtagtt ctgttggtac atattacaca ttcagcaaaa    1560
caactgaaac tcaactgaac ttgtttatac tttgacacag ggtctagcaa aggaaacaac    1620
aatgggaggt agaggtcgtg tggcaaagtg gaagttcaag ggaagaagcc tctctcaagg    1680
gttccaaaca caaagccacc attcactgtt ggccaactca agaaagcaat tccaccacac    1740
tgctttcagc gctccctcct cacttcattc tcctatgttg tttatgacct ttcatttgcc    1800
ttcattttct acattgccac cacctacttc cacctccttc ctcaaccctt ttccctcatt    1860
gcatggccaa tctattgggt tctccaaggt tgccttctca ctggtgtgtg ggtgattgct    1920
cacgagtgtg gtcaccatgc cttcagcaag taccaatggg ttgatgatgt tgtgggtttg    1980
acccttcact caacactttt agtcccttat ttctcatgga aaataagcca tcgccgccat    2040
cactccaaca caggttccct tgaccgtgat gaagtgtttg tcccaaaacc aaaatccaaa    2100
gttgcatggt tttccaagta cttaaacaac cctctaggaa gggctgtttc tcttctcgtc    2160
acactcacaa tagggtggcc tatgtattta gccttcaatg tctctggtag accctatgat    2220
agttttgcaa gccactacca cccttatgct cccatatatt ctaaccgtga gaggcttctg    2280
atctatgtct ctgatgttgc tttgttttct gtgacttact ctctctaccg tgttgcaacc    2340
ctgaaagggt tggtttggct gctatgtgtt tatggggtgc ctttgctcat tgtgaacggt    2400
tttcttgtga ctatcacata tttgcagcac acacactttg ccttgcctca ttacgattca    2460
tcagaatggg actggtgga gggagctttg gcaactacgg acagagatta tgggattctg    2520
aacaaggtgt ttcatcacat aactgatact catgtggctc accatctctt ctctacaatg    2580
ccacattacc atgcaatgga ggcaaccaat gcaatcaagc caatattggg tgagtactac    2640
caatttgatg acacaccatt ttacaaggca ctgtggagag aagcgagaga gtgcctctat    2700
gtggagccag atgaaggaac atccgagaag ggcgtgtatt ggtacaggaa caagtattga    2760
tggagcaacc aatgggccat agtgggagtt atggaagttt tgtcatgtat tagtacataa    2820
ttagtagaat gttataaata agtggatttg ccgcgtaaatg actttgtgtg tattgtgaaa    2880
cagcttgttg cgatcatggt tataatgtaa aaataattct ggtattaatt acatgtggaa    2940
agtgttctgc ttatagcttt ctgcctaaaa tgcacgctgc acgggacaat atcattggta    3000
attttttaa aatctgaatt gaggctactc ataatactat ccataggaca tcaaagacat    3060
gttgcattga ctttaagcag aggttcatct agaggattac tgcataggct tgaactacaa    3120
gtaatttaag ggacgagagc aacttttagct ctaccacgtc gtttacaag gttattaaaa    3180
tcaaattgat cttattaaaa ctgaaaattt gtaataaaat gctattgaaa aattaaaata    3240
tagcaaacac ctaaattgga ctgattttta gattcaaatt taataattaa tctaaattaa    3300
acttaaattt tataatatat gtcttgtaat atatcaagtt ttttttttta ttattgagtt    3360
tggaaacata taataaggaa cattagttaa tattgataat ccactaagat cgacttagta    3420
ttacagtatt tggatgattt gtatgagata ttcaaacttc actcttatca taatagagac    3480
aaaagttaat actgatggtg gagaaaaaaa aatgttattg ggagcatatg gtaagataag    3540
acggataaaa atatgctgca gcctggagag ctaatgtatt tttggtgaa gttttcaagt    3600
gacaactatt catgatgaga acacaataat attttctact tacctatccc acataaaata    3660
ctgattttaa taatgatgat aaataatgat taaaatattt gattcttgt taagagaaat    3720
aaggaaaca taaatattct catggaaaaa tcagcttgta ggagtagaaa ctttctgatt    3780
ataatttaa tcaagtttaa ttcattcttt taattttatt attagtacaa aatcattctc    3840
ttgaatttag agatgtatgt tgtagcttaa tgtaattttt ttattttat aataaaattc    3900
aagcagtcaa atttcatcca aataatcgtg ttcgtgggtg taagtcagtt attccttctt    3960
atcttaatat acacgcaaag gaaaaaataa aaataaaatt cgaggaagcg cagcagcagc    4020
tgataccacg ttggttgacg aaactgataa aaagcgctgt cattgtgtct ttgtttgatc    4080
atcttcacaa tcacatctcc agaacacaaa gaagagtgac ccttcttctt gttattccac    4140
ttgcgttagg tttctacttt cttctctctc tctctctctc tcttcattcc tcatttttcc    4200
ctcaaacaat caatcaattt tcattcagat tcgtaaattt ctcgattaga tcacgggggtt    4260
aggtctccca ctttatcttt tcccaagcct ttctcttttcc ccctttccct gtctgcccca    4320
taaaattcag gatcggaaac gaactgggtt cttgaatttc actctagatt ttgacaaatt    4380
cgaagtgtgc atgcactgat gcgacccact ccccttttt tgcattaaac aattatgaat    4440
tgaggttttt cttgcgatca tcattgcttg aattgaatca tattaggttt agattct      4497
```

<210> 2
<211> 420
<212> DNA
<213> Glycine max

```
<400> 2
gtaaattaaa ttgtgcctgc acctcgggat atttcatgtg gggttcatca tatttgttga      60
ggaaaagaaa ctcccgaaat tgaattatgc atttatatat ccttttttcat ttctagattt     120
cctgaaggct taggtgtagg cacctagcta gtagctacaa tatcagcact tctctctatt      180
gataaacaat tggctgtaat gccgcagtag aggacgatca caacatttcg tgctggttac      240
tttttgtttt atggtcatga tttcactctc tctaatctct ccattcattt tgtagttgtc      300
attatctttta gattttttcac tacctggttt aaaattgagg gattgtagtt ctgttggtac    360
atattacaca ttcagcaaaa caactgaaac tcaactgaac ttgtttatac tttgacacag      420
```

```
<210> 3
<211> 4010
<212> DNA
<213> Glycine max
```

```
<400> 3
acaaagcctt tagcctatgc tgccaataat ggataccaac aaaagggttc ttctttttgat      60
tttgatccta gcgctcctcc accgtttaag attgcagaaa tcagagcttc aataccaaaa      120
cattgctggg tcaagaatcc atggagatcc ctcagttatg ttctcaggga tgtgcttgta      180
attgctgcat tggtggctgc agcaattcac ttcgacaact ggcttctctg ctaatctat       240
tgccccattc aaggcacaat gttctgggct ctctttgttc ttggacatga ttggtaataa      300
ttttttgtgtt tcttactctt ttttttttttt ttttgtttat gatatgaatc tcacacattg    360
ttctgttatg tcatttcttc ttcatttggc tttagacaac ttaaatttga gatcttatt       420
atgttttttgc ttatatggta aagtgattct tcattattttc attcttcatt gattgaattg     480
aacagtggcc atggaagctt ttcagatagc cctttgctga atagcctggt gggacacatc       540
ttgcattcct caattcttgt gccataccat ggatggttag ttcatactgg cttttttgtt        600
tgttcatttg tcattgaaaa aaaatctttt gttgattcaa ttatttttat agtgtgtttg        660
gaagcccgtt tgagaaaata agaaatcgca tctggaatgt gaaagttata actatttagc        720
ttcatctgtc gttgcaagtt cttttattgg ttaaattttt atagcgtgct aggaaaccca       780
ttcgagaaaa taagaaatca catctggaat gtgaaagtta taactgttag cttctgagta       840
aacgtggaaa aaccacattt tggatttgga accaaatttt atttgataaa tgacaaccaa       900
attgattttg atggattttg caggagaatt agccacagaa ctcaccatga aaaccatgga       960
cacattgaga aggatgagtc atgggttcca gtatgtgatt aattgcttct cctatagttg      1020
ttcttgattc aattacattt tatttatttg gtagtgccaa gaaaaaaggg aatctttatg      1080
cttcctgagg ctgttcttga acatggctct tttttatgtg tcattatctt agttaacaga      1140
gaagatttac aagaatctag acagcatgac aagactcatt agattcactg tgccatttcc      1200
atgtttgtgt atccaatta tttggtgagt gattttttga cttggaagac aacaacacat       1260
tattattata atatggttca aaacaatgac ttttttcttta tgatgtgaac tccatttttt     1320
agttttcaag aagccccgga aaggaaggct ctcacttcaa tccctacagc aatctgtttc      1380
cacccagtga gagaaaagga atagcaatat caacactgtg ttgggctacc atgtttttctt      1440
tgcttatcta tctctcattc attaactagt ccacttctag tgctcaagct ctatggaatt       1500
ccatattggg taactaaatt actcctacat tgttacttt tcctccttt ttttattatt         1560
tcaattctcc aattggaaat ttgaaatagt taccataatt atgtaattgt ttgatcatgt      1620
gcagatgttt gttatgtggc tggactttgt cacatacttg catcaccatg gtcaccacca      1680
gaaactgcct tggtaccgcg gcaaggtaac aaaaataaat agaaatagt gggtgaacac       1740
ttaaatgcga gatagtaata cctaaaaaaa gaaaaaaata taggtataat aaataatata      1800
actttcaaaa taaaaagaaa tcatagagtc tagcgtagtg tttggagtga aatgatgttc      1860
acctaccatt actcaaagat tttgttgtgt cccttagttc attcttatta ttttacatat      1920
cttacttgaa aagacttttt aattattcat tgagatctta aagtgactgt taaattaaaa      1980
taaaaacaa gtttgttaaa acttcaaata aataagagtg aagggagtgt catttgtctt       2040
ctttctttta ttgcgttatt aatcacgttt ctcttctctt tttttttttt cttctctgct     2100
ttccacccat tatcaagttc atgtgaagca gtggcggatc tatgtaaatg agtgggggggc      2160
aattgcaccc acaagatttt attttttatt tgtacaggaa taataaaata aaactttgcc      2220
cccataaaaa ataaatattt tttcttaaaa taatgcaaaa taaatataag aaataaaaag      2280
agaataaatt attattaatt ttattatttt gtacttttta tttagttttt ttagcggtta      2340
```

- 3 -

```
gatttttttt tcatgacatt atgtaatctt ttaaaagcat gtaatatttt tattttgtga   2400
aaataaatat aaatgatcat attagtctca gaatgtataa actaataata attttatcac   2460
taaaagaaat tctaatttag tccataaata agtaaaacaa gtgacaatta tattttatat   2520
ttacttaatg tgaaataata cttgaacatt ataataaaac ttaatgacag gagatattac   2580
atagtgccat aaagatattt taaaaaataa aatcattaat acactgtact actatataat   2640
attcgatata tattttaac atgattctca atagaaaaat tgtattgatt atatttatt    2700
agacatgaat ttacaagccc cgtttttcat ttatagctct tacctgtgat ctattgtttt   2760
gcttcgctgt ttttgttggt caagggactt agatgtcaca atattaatac tagaagtaaa   2820
tatttatgaa aacatgtacc ttacctcaac aaagaaagtg tggtaagtgg caacacacgt   2880
gttgcatttt tggcccagca ataacacgtg ttttgtggt gtactaaaat ggacaggaat    2940
ggagttattt aagaggtggc ctcaccactg tggatcgtga ctatggttgg atcaataaca   3000
ttcaccatga cattggcacc catgttatcc accatctttt cccccaaatt cctcattatc   3060
acctcgttga agcggtacat tttattgctt attcacctaa aaacaataca attagtacat   3120
ttgttttatc tcttggaagt tagtcatttt cagttgcatg attctaatgc tctctccatt   3180
cttaaatcat gtttcacac ccacttcatt taaaataaga acgtgggtgt tattttaatt    3240
tctattcact aacatgagaa attaacttat ttcaagtaat aattttaaaa tattttttatg   3300
ctattatttt attacaaata attatgtata ttaagtttat tgattttata ataattatat   3360
taaaattata tcgatattaa tttttgattc actgatagtg ttttatattg ttagtactgt   3420
gcatttattt taaaattggc ataaataata tatgtaacca gctcactata ctatactggg   3480
agcttggtgg tgaaaggggt tcccaaccct cctttctagg tgtacatgct ttgatacttc   3540
tggtaccttc ttatatcaat ataaattata ttttgctgat aaaaaaacat ggttaaccat   3600
taaattcttt ttttaaaaaa aaaactgtat ctaaactttg tattattaaa aagaagtctg   3660
agattaacaa taaactaaca ctcatttgga ttcactgcag acacaagcag caaaaccagt   3720
tcttggagat tactaccgtg agccagaaag atctgcgcca ttaccatttc atctaataaa   3780
gtatttaatt cagagtatga gacaagacca cttcgtaagt gacactggag atgttgttta   3840
ttatcagact gattctctgc tcctccactc gcaacgagac tgagtttcaa acttttgggg   3900
ttattattta ttgattctag ctactcaaat tactttttt ttaatgttat gttttttgga    3960
gtttaacgtt ttctgaacaa cttgcaaatt acttgcatag agagacatgg               4010
```

```
<210> 4
<211> 192
<212> DNA
<213> Glycine max

<400> 4
gtaataattt ttgtgtttct tactcttttt ttttttttt tgtttatgat atgaatctca   60
cacattgttc tgttatgtca tttcttcttc atttggcttt agacaactta aatttgagat   120
ctttattatg ttttttgctta tatggtaaag tgattcttca ttatttcatt cttcattgat   180
tgaattgaac ag                                                          192
```

```
<210> 5
<211> 348
<212> DNA
<213> Glycine max

<400> 5
gttagttcat actggctttt ttgtttgttc atttgtcatt gaaaaaaaat cttttgttga   60
ttcaattatt tttatagtgt gtttggaagc ccgtttgaga aaataagaaa tcgcatctgg   120
aatgtgaaag ttataactat ttagcttcat ctgtcgttgc aagttctttt attggttaaa   180
ttttttatagc gtgctaggaa acccattcga gaaaataaga aatcacatct ggaatgtgaa   240
agttataact gttagcttct gagtaaacgt ggaaaaacca cattttggat ttggaaccaa   300
attttatttg ataaatgaca accaaattga ttttgatgga ttttgcag                  348
```

```
<210> 6
<211> 142
```

```
<212> DNA
<213> Glycine max

<400> 6
gtatgtgatt aattgcttct cctatagttg ttcttgattc aattacattt tatttatttg    60
gtaggtccaa gaaaaaaggg aatctttatg cttcctgagg ctgttcttga acatggctct   120
tttttatgtg tcattatctt ag                                            142

<210> 7
<211> 1231
<212> DNA
<213> Glycine max

<400> 7
gtaacaaaaa taaatagaaa atagtgggtg aacacttaaa tgcgagatag taatacctaa    60
aaaaagaaaa aaatataggt ataataaata atataacttt caaaataaaa agaaatcata   120
gagtctagcg tagtgtttgg agtgaaatga tgttcaccta ccattactca aagatttttgt  180
tgtgtcccct agttcattct tattatttta catatcttac ttgaaaagac tttttaatta   240
ttcattgaga tcttaaagtg actgttaaat taaaataaaa aacaagtttg ttaaaacttc   300
aaataaataa gagtgaaggg agtgtcattt gtcttctttc ttttattgcg ttattaatca   360
cgtttctctt ctcttttttt tttttcttct ctgctttcca cccattatca agttcatgtg   420
aagcagtggc ggatctatgt aaatgagtgg ggggcaattg cacccacaag attttatttt   480
ttatttgtac aggaataata aaataaaact ttgcccccat aaaaaataaa tattttttct   540
taaaataatg caaaataaat ataagaaata aaaagagaat aaattattat taattttatt   600
attttgtact ttttatttag ttttttttagc ggttagattt ttttttcatg acattatgta   660
atctttttaaa agcatgtaat attttttattt tgtgaaaata aatataaatg atcatattag   720
tctcagaatg tataaactaa taataatttt atcactaaaa gaaattctaa tttagtccat   780
aaataagtaa aacaagtgac aattatattt tatatttact taatgtgaaa taatacttga   840
acattataat aaaacttaat gacaggagat attacatagt gccataaaga tattttaaaa   900
aataaaatca ttaatacact gtactactat ataatattcg atatatattt ttaacatgat   960
tctcaataga aaaattgtat tgattatatt ttattagaca tgaatttaca agccccgttt  1020
ttcatttata gctcttacct gtgatctatt gttttgcttc gctgttttttg ttggtcaagg  1080
gacttagatg tcacaatatt aatactagaa gtaaatattt atgaaacat gtaccttacc    1140
tcaacaaaga aagtgtggta agtggcaaca cacgtgttgc atttttggcc cagcaataac  1200
acgtgttttt gtggtgtact aaaatggaca g                                  1231

<210> 8
<211> 626
<212> DNA
<213> Glycine max

<400> 8
gtacatttta ttgcttattc acctaaaaac aatacaatta gtacatttgt tttatctctt    60
ggaagttagt cattttcagt tgcatgattc taatgctctc tccattctta aatcatgttt   120
tcacacccac ttcatttaaa ataagaacgt gggtgttatt ttaatttcta ttcactaaca   180
tgagaaatta acttatttca agtaataatt ttaaaatatt tttatgctat tatttttatta   240
caaataatta tgtatattaa gtttattgat tttataataa ttatattaaa attatatcga   300
tattaattttt tgattcactg atagtgtttt atattgttag tactgtgcat ttattttaaa   360
attggcataa ataatatatg taaccagctc actatactat actgggagct tggtggtgaa   420
aggggttccc aaccctcctt tctaggtgta catgctttga tacttctggt accttcttat   480
atcaatataa attatatttt gctgataaaa aaacatggtt aaccattaaa ttctttttttt   540
aaaaaaaaaa ctgtatctaa actttgtatt attaaaaaga agtctgagat taacaataaa   600
ctaacactca tttggattca ctgcag                                        626

<210> 9
```

```
<211> 18
<212> DNA
<213> SyntheticArtificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 9
atacaagcca ctaggcat                                                18

<210> 10
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 10
gattggccat gcaatgaggg aaaagg                                       26

<210> 11
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 11
cuacuacuac uactcgagac aaagccttta gcctatg                           37

<210> 12
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 12
caucaucauc auggatccca tgtctctcta tgcaag·                           36

<210> 13
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 13
acgaattcct cgaggtaaat taaattgtgc ctgc                              34

<210> 14
```

```
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 14
gcgagatcta tcgatctgtg tcaaagtata aac                    33

<210> 15
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 15
gatcgatgcc cggggtaata atttttgtgt                        30

<210> 16
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 16
cacgcctcga gtgttcaatt caatcaatg                         29

<210> 17
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 17
cactcgagtt agttcatact ggct                              24

<210> 18
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 18
cgcatcgatt gcaaaatcca tcaaa                             25

<210> 19
```

```
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 19
cuacuacuac uactcgagcg taaatagtgg gtgaacac                              38

<210> 20
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 20
caucaucauc auctcgagga attcgtccat tttagtacac c                         41

<210> 21
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 21
cuacuacuac uactcgaggc gcgtacattt tattgctta                            39

<210> 22
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 22
caucaucauc auctcgagga attctgcagt gaatccaaat g                         41

<210> 23
<211> 1734
<212> DNA
<213> Glycine max

<400> 23
actatagggc acgcgtggtc gacggcccgg gctggtcctc ggtgtgactc agccccaagt     60
gacgccaacc aaacgcgtcc taactaaggt gtagaagaaa cagatagtat ataagtatac    120
catataagag gagagtgagt ggagaagcac ttctcctttt tttttctctg ttgaaattga    180
aagtgttttc cgggaaataa ataaaataaa ttaaaatctt acacactcta ggtaggtact    240
tctaatttaa tccacacttt gactctatat atgttttaaa aataattata atgcgtactt    300
acttcctcat tatactaaat ttaacatcga tgattttatt ttctgtttct cttctttcca    360
```

- 8 -

```
cctacataca tcccaaaatt tagggtgcaa ttttaagttt attaacacat gttttagct       420
gcatgctgcc tttgtgtgtg ctcaccaaat tgcattcttc tctttatatg ttgtatttga       480
attttcacac catatgtaaa caagattacg tacgtgtcca tgatcaaata caaatgctgt       540
cttatactgg caatttgata aacagccgtc cattttttct ttttctcttt aactatatat       600
gctctagaat ctctgaagat tcctctgcca tcgaatttct ttcttggtaa caacgtcgtc       660
gttatgttat tattttattc tattttttatt ttatcatata tatttcttat tttgttcgaa      720
gtatgtcata ttttgatcgt gacaattaga ttgtcatgta ggagtaggaa tatcacttta       780
aaacattgat tagtctgtag gcaatattgt cttctttttc ctcctttatt aatatatttt       840
gtcgaagttt taccacaagg ttgattcgct ttttttgtcc ctttctcttg ttctttttac       900
ctcaggtatt ttagtctttc atggattata agatcactga gaagtgtatg catgtaatac       960
taagcaccat agctgttctg cttgaattta tttgtgtgta aattgtaatg tttcagcgtt      1020
ggctttccct gtagctgcta caatggtact gtatatctat tttttgcatt gttttcattt      1080
tttcttttac ttaatcttca ttgctttgaa attaataaaa caatataata tagtttgaac      1140
tttgaactat tgcctattca tgtaattaac ttattcactg actcttattg ttttctggt       1200
agaattcatt ttaaattgaa ggataaatta agaggcaata cttgtaaatt gacctgtcat      1260
aattacacag gaccctgttt tgtgcctttt tgtctctgtc tttggtttttg catgttagcc     1320
tcacacagat atttagtagt tgttctgcat acaagcctca cacgtatact aaaccagtgg      1380
acctcaaagt catggcctta cacctattgc atgcgagtct gtgacacaac ccctggtttc      1440
catattgcaa tgtgctacgc cgtcgtcctt gtttgtttcc atatgtatat tgataccatc      1500
aaattattat atcatttata tggtctggac cattacgtgt actctttatg acatgtaatt      1560
gagtttttta attaaaaaaa tcaatgaaat ttaactacgt agcatcatat agagataatt      1620
gactagaaat ttgatgactt attctttcct aatcatattt tcttgtattg atagccccgc      1680
tgtcccttttt aaactcccga gagagtataa aactgcatcg aatattacaa gatg           1734
```

```
<210> 24
<211> 405
<212> DNA
<213> Glycine max
```

```
<400> 24
gtatgatgct aaattaaatt gtgcctgcac cccaggatat ttcatgtggg attcatcatt        60
tattgaggaa aactctccaa attgaatcgt gcatttatat ttttttttcca tttctagatt      120
tcttgaaggc ttatggtata ggcacctaca attatcagca cttctctcta ttgataaaca       180
attggctgta ataccacagt agagaacgat cacaacattt tgtgctggtt accttttgtt       240
ttatggtcat gatttcactc tctctaatct gtcacttccc tccattcatt ttgtacttct       300
catatttttc acttcctggt tgaaaattgt agttctcttg gtacatacta gtattagaca       360
ttcagcaaca acaactgaac tgaacttctt tatactttga cacag                       405
```

```
<210> 25
<211> 98
<212> DNA
<213> Glycine max
```

```
<400> 25
gtgagtgatt ttttgacttg gaagacaaca acacattatt attataatat ggttcaaaac        60
aatgactttt tctttatgat gtgaactcca tttttttag                              98
```

```
<210> 26
<211> 115
<212> DNA
<213> Glycine max
```

```
<400> 26
gtaactaaat tactcctaca ttgttacttt ttcctccttt tttttattat ttcaattctc        60
caattggaaa tttgaaatag ttaccataat tatgtaattg tttgatcatg tgcag           115
```

```
<210> 27
<211> 778
<212> DNA
<213> Glycine max

<220>
<221> misc_feature
<222> (1)...(778)
<223> n = A,T,C or G

<400> 27
atacaagcca ctaggcatgg taaattaaat tgtgcctgca cctcgggata tttcatgtgg          60
ggttcatcat atttgttgag gaaaagaaac tcccgaaatt gaattatgca tttatatatc         120
ctttttcatt tctagatttc ctgaaggctt aggtgtaggc acctagctag tagctacaat         180
atcagcactt ctctctattg ataaacaatt ggctgtaatg ccgcagtaga ggacgatcac         240
aacatttcgt gctggttact ttttgtttta tggtcatgat ttcactctct ctaatctctc         300
cattcatttt gtagttgtca ttatctttag attttttcact acctggtttta aaattgaggg        360
attgtagttc tgttggtaca tattacacat tcagcaaaac aactgaaact caactgaact          420
tgtttatact ttgacacagg gtctagcaaa ggaaacaaca atgggaggta gaggtcgtgt          480
ggccaaagtg gaagttcaag ggaagaagcc tctctcaagg gttccaaaca caaagccacc          540
attcactgtt ggccaactca agaaagcaat tccaccacac tgctttcagc gctccctcct          600
cacttcattc tcctatgttg tttatgacct ttcatttgcc ttcattttct acattgccac          660
cacctacttc cacctccttc ctcaacccctt ttccctcatt gcatggccaa tcaagccgaa         720
ttctgcagat atccatcaca tggcggcggn tggngnaggn ntntanaggg cccaattc           778

<210> 28
<211> 148
<212> DNA
<213> Glycine max

<400> 28
gtaatctcac tctcacactt tctttataca tcgcacacca gtgtgggtta tttgcaacct          60
acaccgaagt aatgccctat aattaatggg gttaacacat gtccaagtcc aatattttgt         120
tcacttattt gaacttgaac atgtgtag                                            148

<210> 29
<211> 361
<212> DNA
<213> Glycine max

<400> 29
gtatcccatt taacacaatt tgtttcatta acattttaag agaatttttt tttcaaaata          60
gttttcgaaa ttaagcaaat accaagcaaa ttgttagatc tacgcttgta cttgttttaa         120
agtcaaattc atgaccaaat tgtcctcaca agtccaaacc gtccactatt ttatttcac          180
ctactttata gcccaatttg tcatttggtt acttcagaaa agagaacccc atttgtagta         240
aatatattat ttatgaatta tggtagtttc aacataaaac atatttatgt gcagttttgc         300
catccttcaa aagaagatag aaacttactc catgttactc tgtctatatg taatttcaca         360
g                                                                         361

<210> 30
<211> 36
<212> DNA
<213> Glycine max
```

```
<400> 30
gcgatcgatg tatgatgcta aattaaattg tgcctg                              36

<210> 31
<211> 30
<212> DNA
<213> Glycine max

<400> 31
gcggaattcc tgtgtcaaag tataaagaag                                     30

<210> 32
<211> 19
<212> DNA
<213> Glycine max

<400> 32
catgctttct gtgcttctc                                                 19

<210> 33
<211> 19
<212> DNA
<213> Glycine max

<400> 33
gttgatccaa ccatagtcg                                                 19
```

BLANK PAGE